Hey Alexa, is this Skill Safe?
Taking a Closer Look at the Alexa Skill Ecosystem

Christopher Lentzsch¹, Sheel Jayesh Shah², Benjamin Andow³, Martin Degeling¹, Anupam Das², William Enck²

¹ Ruhr-Universität Bochum
² North Carolina State University
³ Google Inc.

LASER Workshop, 2021
Voice Assistants

- Voice-based user interface
- 3+ billion devices
- Placed in personal settings
- New security/privacy risks
How Alexa Works?

Alexa, how is weather today?

User → Echo Device → Amazon Skill Service → Alexa weather skill
How Alexa Works?

It is bright and sunny today!

User ➔ Reads Response ➔ Processes Response ➔ Response ➔ Resolves Intent
Skills: Applications for Voice Assistants

- Enhance basic functionalities
- Integrate with third parties
- Instant activation
  - Alexa, open `<invocation name>`
- Vetted prior to publication
Research Questions

1. What limitations exist in the current skill vetting process?
   - Duplicate invocation names
   - Developer names
   - Registered intents
   - Permission model

2. How effective are skill squatting attacks?
   - Which patterns are more effective?

3. Is the requirement of providing a privacy policy effective?
   - Are data types properly disclosed
Our Data Collection Methodology

- Crawled the top 7 skill stores for meta data
- Rented servers in 5 different locations (DigitalOcean)
  - US, Canada, UK, Germany (for DE, FR), Singapore (for AU, JP)
  - To prevent geo-blocking
- Crawler used SELENIUM
  - Firefox Browser, headful, less likely detected as bot
Gathering Skill IDs

- Traversed each skill category separately
  - 23 categories (in Jan 2020)
  - Max 400 pages per category
- Track skills through unique IDs (ASIN), e.g. B07KX2CSXM
- Also downloaded skills appearing in the recommended list
Extracting Skill Metadata

- Downloaded metadata from skill information page
  - Parsed using beautiful soup
  - Downloaded privacy policies (if available)
- Took around 9 days for the US Skill store
  - At time our crawler was blocked and had to wait
Example of Metadata
Brief Summary of Data

- Total 90,194 unique skills

In the US store there are over 35k unique skills with over 13k unique developers
Question: How does Alexa select skills?

- Many duplicate skill invocation names
  - For *space facts* there are 81 skills
  - Skills are auto enabled
Methodology to Test Skill Selection Process

- Find skills that have same invocation name
  - Differ in other observable attributes
- Test invocation names
  - Alexa, open <invocation name>
  - Rerun test three times
- Analyze results
Example: Skills with Same Invocations

<table>
<thead>
<tr>
<th>Invocation</th>
<th>UID</th>
<th>Name</th>
<th>Developer</th>
<th># Ratings</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>grandfather clock</td>
<td>B071DWVZQW</td>
<td>Sleep Sounds: Grandfather</td>
<td>Voice Apps, LLC.</td>
<td>115.0</td>
<td>4.5 out of 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grandfather clock</td>
<td>B076452X2P</td>
<td>Grandfather Clock</td>
<td>ut666</td>
<td>57.0</td>
<td>3.8 out of 5</td>
</tr>
</tbody>
</table>

- Example: `<grandfather clock>`
- Different number of ratings
- Different avg. rating
What Factors to Test?

- Available metadata
  - Number of ratings
  - Avg. rating
  - Age of skill
  - Category
  - Dynamic (true/false)
  - Guidance (true/false)
  - Permissions
  - Privacy policy (present/none)
  - Terms of service (present/none)

Note: not all properties provided enough statistically significant samples
Automate Testing using Amazon Polly

• TTS-Service Amazon Polly
  ▪ 'Salli', female, en-US voice
  ▪ "Alexa, start unicorn facts"
Create User Accounts for Testing

• Tested across three new independent accounts

• Created Accounts with US-bound IPs (VPN-Service)
  ▪ To avoid localization, e.g., german language or store
  ▪ To avoid rate limiting
Example: Testing

- Play invocation
  - "Alexa, start unicorn facts"

- Wait 5 seconds
  - "Alexa, stop"

- Wait 3 seconds
  - "Alexa, exit"

- Wait 3 seconds

- Get Activity-Log
  - Curl-Call to REST-API
{  "cards": [  {   "cardType": "A2SEnableSkillCard",  
"developerName": "Envy Eden",
"examplePhrase": "Alexa, open Unicorn Facts",
"hint": null,
"originIntentType": "LaunchNativeAppIntent",
"playbackAudioAction": {   "actionType": "PlayAudioAction",
"mainText": "Alexa heard: \"alexa start unicorn facts\""},
"primaryActions": [   {    "actionType": "NavigateAction",
"mainText": "View Skill Details",
"route": "skills/dp/B07DL8X97K",
"routeAddOnComponent": null,
"serviceName": null,
"subText": "Description, additional phrases, reviews, developer terms of use, privacy policy, and other details",
"subTextRoute": null  
}  ],
"skillName": "Unicorn Facts",
}  ]}
Fisher’s Exact Test  
for # of ratings

<table>
<thead>
<tr>
<th></th>
<th>more ratings</th>
<th>less ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>activated</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>not activated</td>
<td>10</td>
<td>40</td>
</tr>
</tbody>
</table>

• Odds ratio = 16.0
• p-value <0.0001
## Other Factors

<table>
<thead>
<tr>
<th>Attribute 1</th>
<th>Attribute 2</th>
<th>Favored Attribute</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different number of ratings</td>
<td>more ratings</td>
<td></td>
<td>&lt; 0.0001 ****</td>
</tr>
<tr>
<td>Different avg. rating</td>
<td>higher avg. rating</td>
<td></td>
<td>0.00012 ***</td>
</tr>
<tr>
<td>Age of skill</td>
<td></td>
<td></td>
<td>0.84162</td>
</tr>
<tr>
<td>Content advisory</td>
<td></td>
<td></td>
<td>0.54874</td>
</tr>
<tr>
<td>Same number of ratings</td>
<td>Different avg. rating</td>
<td>higher avg. rating</td>
<td>0.03476 *</td>
</tr>
<tr>
<td>Same number of ratings</td>
<td>Age of skill</td>
<td></td>
<td>0.31734</td>
</tr>
<tr>
<td>Same number of ratings</td>
<td>Content advisory</td>
<td></td>
<td>0.84161</td>
</tr>
</tbody>
</table>
Manipulating Attributes

• We launched our own skill pair
  ▪ Identical invocation name
  ▪ Increased num. of ratings and interaction
    – Failed to change skill selection

Takeaway: Positive correlation with rating, but it does not necessarily imply causation.
Question: Can I use any Developer Name?

• Skill page shows a developer name
• Published our skills as:
  ▪ Microsoft, Ring, Samsung and Withings
• Philips got rejected

Takeaway: An attacker can getaways with publishing skills using well-known company names.
Question: Can I Register Dormant Intents?

- Skills have intents
  - Each intent can many slots (i.e., data type)
Code Change after Approval

- Generate intent model (dormant intent present)
- Skill is vetted and approved
- Backend code is changed
- User is directed to trigger dormant intent
- Adversary gets user’s phone number
**Takeaway:** An adversary can change the backend code after approval to coax users into revealing sensitive information.
Question: Do Skills Bypass Permission?

Extract skill description

Use regular expressions to search for permission protected data types

Manually vet and activate skill for verification

Skill description define what the skill does

Phone number, location, e-mail, name

Ignore account linking skills and skills requesting permissions
## Regex Used

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Regular Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>`\b(your)\s+(((whole</td>
</tr>
<tr>
<td>Location</td>
<td>`\b(your)\s+((home</td>
</tr>
<tr>
<td>Phone Number</td>
<td>`\b(your)\s+((home</td>
</tr>
<tr>
<td>Email</td>
<td>`\b(your)\s+((home</td>
</tr>
</tbody>
</table>
Some Skills ask Sensitive Data

<table>
<thead>
<tr>
<th>Filtering mechanism</th>
<th>Data Type</th>
<th>Unique skills</th>
<th>w/o PP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skills detected through regular expression</td>
<td>Name: 432</td>
<td>Email: 417</td>
<td>Phone: 242</td>
</tr>
<tr>
<td>After manually inspecting skill description</td>
<td>Name: 109</td>
<td>Email: 26</td>
<td>Phone: 108</td>
</tr>
<tr>
<td>Verbally request data</td>
<td>Name: 65</td>
<td>Email: 4</td>
<td>Phone: 33</td>
</tr>
<tr>
<td>Non-verbally request data</td>
<td>Name: 1</td>
<td>Email: 1</td>
<td>Phone: 1</td>
</tr>
<tr>
<td>Does not request data</td>
<td>Name: 20</td>
<td>Email: 7</td>
<td>Phone: 4</td>
</tr>
<tr>
<td>Skill invocable but non-functional</td>
<td>Name: 19</td>
<td>Email: 12</td>
<td>Phone: 62</td>
</tr>
<tr>
<td>Skill not available in store</td>
<td>Name: 4</td>
<td>Email: 2</td>
<td>Phone: 8</td>
</tr>
</tbody>
</table>

* Some skills access multiple data types, hence the summation across different data types will be slightly higher than the number of unique skills.
An example of True Positive

Developer Quotes
by Tony
Rated: Guidance Suggested
Free to Enable

“Alexa open geek quotes”  “tell me a quote”

Shown in: English (US)  See all supported languages

Description
A fun app that tells geeky computer jokes. When you start up it tries to get to know you by asking you your name and favorite computer language, then it tells you geeky computer jokes.

Get this Skill
Enable

By enabling, this skill can be accessed on all your available Alexa devices.
An example of False Positive

Presidential Quest
by Will Mundy

Free to Enable

“Alexa, launch Presidential Quest”
“Alexa, ask President... start quest”

Shown in: English (US) See all supported languages

Description
Think you know our presidents? Think again.

Imagine this: your name is Thomas Jefferson, and you've just received word that Aaron Burr is planning to create a new country within your borders. What do you do? Do you sit back and watch as your former Vice-President takes your land? Or do you fight for your country by answering his Presidential Trivia?
Question: Do Skills Disclose Data Practices?

Extract data flows from privacy policy using **PoliCheck (NLP tool)**

Re-map data flows to permission requests

Classify each data flow

Extract statements regarding the collection and share of data

Manually adapt the data type ontology of PoliCheck to match Alexa permissions

Three groups: consistent, partial, inconsistent
Data Practice Disclosure in Privacy Policy

- Are requested permissions covered?
  - Analyzed 1,124 skills with 1,447 permission requests

**Takeaway:** Sensitive data types are not fully disclosed by ~23% of the skills‘ privacy policies.
Skill: “Find me Breakfast” (B07K9NQX5B)

Find me Breakfast
by hermanj13
Rated: Guidance Suggested
1
Free to Enable

“Alexa, open Find me Breakfast.”

“What is the closest place to me?"

Get this Skill
Enable

This skill needs permission to access:
- Device Address

By enabling, this skill can be accessed on all your available Alexa devices.
Privacy Policy of “Find me Breakfast”  
(B07K9NQX5B)

General

When you use our skills you have to talk to Alexa. This voice input is sent to Amazon and us where we use it to understand what our skill should do for you. This is absolutely necessary for our service to give you an appropriate answer.

Data

We never collect or share personal data with our skills.

To improve our services we analyze automatically how often utterances are spoken and other analytics. This is done automatically by Amazon in the Amazon Developer Portal.
Skill: “Wear Assistant” (B072KL1S3G)

Wear Assistant
by FluiBex
Rated: Guidance Suggested
249
Free to Enable

“Alexa, open Wear Assistant”
“Alexa, ask Wear Assistant should I dress today York”

Shown in: English (US) See all supported languages

Get this Skill
Enable

This skill needs permission to access:
- Device Address

By enabling, this skill can be accessed on all your available Alexa devices.
Privacy Policy of “Wear Assistant” (B072KL1S3G)

General

When you use our skills you have to talk to Alexa. This voice input is sent to Amazon and us where we use it to understand what our skill should do for you. This is absolutely necessary for our service to give you an appropriate answer.

Data

We never collect or share personal data with our skills.

For some of our skills, after your consent, we may use your address information in order to automatically detect your location and speed up the interaction. Your location data is never collected on our server, but it is used on the fly during the skill response built. In any moment you can revoke the use of your location information to our skills from the Amazon Alexa app.
What did we learn from our methodology?

- Automated skill activation
  - Tried building a chatbot to interact with skills, but failed to process and respond before timeout

- Detecting squatting skills wasn’t as easy we had thought
  - Had to manually go through similar skill pairs after identifying phonetically similar invocations

- Existing NLP techniques were readily applicable
  - PoliCheck
Going Forward

• Develop systems to automatically interact with skills to determine if they access sensitive data

• Conduct user studies to understand people’s perception of how skills work

• Better security indicators for the voice interface
Contact

Responsible Disclosure
Reported our findings to Amazon and have talked with them.

Our data set is public:
https://alexa-skill-analysis.org

Christopher Lentzsch <christopher.lentzsch@rub.de>
Sheel Jayesh Shah <sshah28@ncsu.edu>
Benjamin Andow <andow@google.com>
Martin Degeling <martin.degeling@rub.de>
Anupam Das <anupam.das@ncsu.edu>
William Enck <whenck@ncsu.edu>